

# SPECTRAL STRUCTURE OF SINGULAR SPECTRUM DECOMPOSITION FOR TIME SERIES

Kenji Kume <sup>\*</sup>      Naoko Nose-Togawa <sup>†</sup>

July 28, 2015

## Abstract

Singular spectrum analysis (SSA) is a nonparametric and adaptive spectral decomposition of a time series. The singular value decomposition of the trajectory matrix and the anti-diagonal averaging leads to a time-series decomposition. In this algorithm, a single free parameter, window length  $K$ , is involved which is the FIR filter length for the time series. There are no generally accepted criterion for the proper choice of the window length  $K$ . Moreover, the proper window length depends on the specific problem which we are interested in. Thus, it is important to monitor the spectral structure of the SSA decomposition and its window length dependence in detail for the practical application. In this paper, based on the filtering interpretation of SSA, it is shown that the decomposition of the power spectrum for the original time series is possible with the filters constructed from the eigenvectors of the lagged-covariance matrix. With this, we can obtain insights into the spectral structure of the SSA decomposition and it helps us for the proper choice of the window length in the practical application of SSA.

---

<sup>\*</sup>Department of Physics, Nara Women's University, Nara 630-8506, Japan.

<sup>†</sup>Research Center for Nuclear Physics, Osaka University, Ibaraki 567-0047, Japan

# 1 Introduction

Singular Spectrum Analysis (SSA) is a nonparametric and adaptive decomposition of a time series. In the SSA, the original time series is separated into arbitrary number of additive components which are grouped and are interpreted as the slowly varying trend, oscillatory and noise components, respectively. It has been applied to many different fields from physics and earth sciences to financial and social sciences [2-7].

In the conventional treatment of SSA, the algorithm consists of four steps:

- 1st step: a time series  $\{x_0, x_1, x_2, \dots, x_{N-1}\}$  of length  $N$  is transformed into multidimensional series by constructing the  $L \times K$  trajectory matrix  $X$ . Where  $K$  is the window length which is an arbitrary positive integer satisfying  $K \leq N/2$  and  $L = N - K + 1$ .
- 2nd step: The singular value decomposition of the trajectory matrix

$$X = X_1 + X_2 + \dots + X_K \quad . \quad (1)$$

- 3rd step: so-called grouping procedure where the above set of rank-one matrices  $X_k$  is partitioned into several groups.
- 4th step: reconstruction process where the decomposition of the time series is completed with the anti-diagonal averaging of the grouped matrix  $X^{(j)} = X_{i_1} + X_{i_2} + X_{i_3} + \dots + X_{i_m}$  which transform  $X^{(j)}$  into Hankel form.

The major objective of the SSA algorithm is to separate the original time series into interpretable set of time series. A single parameter, window length  $K$ , is involved in the algorithm and the results of decomposition depend on the choice of this parameter  $K$ . These points are thoroughly discussed from the *separability* point of view [3,4]. It is argued that the relatively large window lengths are preferable to achieve a better (weak) separation, but for the series with a complex structure, too large window length leads to undesirable decomposition or mixture of the series components. Then, how to choose the proper window length  $K$  is an important point for the practical application of SSA but is not easy to formulate in a general way. The choice of the window lengths have been done so far by investigating the behavior of the eigentriples (singular values, and left and right singular vectors of the trajectory matrix) or the w-correlation matrix.

In this paper, based on the filtering interpretation of SSA [1,8,10,12], the decomposition of the power spectrum for the original time series is shown to be possible and this gives us detailed information about the spectral structure of the SSA decomposition. In the filtering interpretation, the SSA algorithm can be seen as the generation of normalized and complete set of adaptive filters and the operation of these filters to the original data. The SSA decomposition is thus interpreted as the soft and the optimal partitioning of the Fourier space through the adaptively constructed filters. By utilizing the completeness property of the eigenvectors of the lagged-covariance matrix, the filters constructed with these eigenvectors are used for the decomposition of the power spectrum of an original time series. The decomposition of the power spectrum shows the detailed structure of the SSA decomposition and, with this, we can obtain useful information about the spectral structure of the SSA decomposition and this helps us to choose the proper window length  $K$  in the practical application.

In the next Sec. 2, we briefly summarize the conventional treatment of the SSA algorithm. We also describe the filtering interpretation of the SSA and it is adopted for the decomposition of the power spectrum of original time series. In Sec. 3, in order to see how the decomposition of the power spectrum works, we have examined two cases: the schematic time series with the superposition of the pure harmonic components and the daily currency exchange rates of EUR/USD. The summary and conclusions are given in Sec. 4.

## 2 Algorithm for the singular spectrum analysis

In this section, we briefly describe the SSA algorithm. In the filtering interpretation, the SSA algorithm consists of two steps. First, optimal and adaptive construction of the normalized and complete set of filters through the eigenvectors of the lagged-covariance matrix. Second step is the operation of these filters to the original time series. In the conventional treatment of the SSA, these two procedures are mixed into the singular value decomposition and the anti-diagonal averaging procedure. Because of the completeness property of the filters generated in SSA, not only the time series but also the power spectrum can be decomposed into arbitrary number of *additive* com-

ponents. In order to show this, we assume the periodic boundary condition for the time series to avoid subtleties coming from the starting or the end points of a time series.

## 2.1 Conventional SSA algorithm

Let us consider a real-valued time series  $\{x_0, x_1, x_2, \dots, x_{N-1}\}$  of sufficient length. First, we define the  $L \times K$  trajectory matrix  $X$  as

$$X_{ij} = x_{i+j-2} \quad (i = 1, 2, \dots, L; j = 1, 2, \dots, K) \quad , \quad (2)$$

where window length  $K$  is the arbitrary positive integer satisfying  $K \leq N/2$  and  $L = N - K + 1$ . Next step is the singular value decomposition of the above trajectory matrix  $X$ . To do this, the lagged-covariance matrix  $C = X^T X$  is introduced and the eigenvalue equation is solved

$$C \mathbf{v}^{(k)} = \lambda_k \mathbf{v}^{(k)} \quad . \quad (3)$$

The eigenvalues  $\lambda_k$  are ordered as the decreasing order of magnitude ( $\lambda_1, \lambda_2, \dots, \lambda_K$ ) and the corresponding eigenvectors are denoted as  $\mathbf{v}^{(k)}$ . We define the  $L$  dimensional vector as  $\mathbf{w}^{(k)} = X \mathbf{v}^{(k)} / \sqrt{\lambda_k}$ . Then the  $L \times K$  rank-one matrices  $X_k$  are expressed as  $X_k = \sqrt{\lambda_k} \mathbf{w}^{(k)} \mathbf{v}^{(k)T}$ . Then the singular value decomposition of the trajectory matrix is expressed as

$$X = \sum_{k=1}^K X_k = \sum_{k=1}^K \sqrt{\lambda_k} \mathbf{w}^{(k)} \mathbf{v}^{(k)T} \quad . \quad (4)$$

The third step is the partitioning of the matrices  $X_k$  into several groups and is called as the grouping procedure. The fourth step is the reconstruction of the time series. For the grouped matrix  $X^{(j)} = X_{i_1} + X_{i_2} + X_{i_3} + \dots + X_{i_m}$ , we carry out the anti-diagonal average which transform  $X^{(j)}$  into Hankel form. Then the time series are reconstructed giving the SSA decomposition.

## 2.2 Filtering interpretation of SSA algorithm

As are discussed in Refs. [8,10-12], above mentioned SSA algorithm can be viewed as the two-step point-symmetric filtering of the time series. The filtering interpretation is much more transparent in Fourier space, and the normalization and the completeness properties of the filters play a crucial role

in SSA decomposition. To see this, we define the discrete Fourier transform as

$$\hat{x}_\alpha = \frac{1}{\sqrt{L}} \sum_{j=0}^{L-1} \exp(2\pi i \alpha j / L) x_j \quad , \quad (5)$$

and

$$\hat{v}_\alpha^{(k)} = \sum_{j=0}^{K-1} \exp(2\pi i \alpha j / L) v_{j+1}^{(k)} \quad , \quad (6)$$

where the components of the vector  $\mathbf{v}^{(k)}$  is denoted as  $(v_1^{(k)}, v_2^{(k)}, \dots, v_K^{(k)})^T$ . In above Eq.(6), we consider the Fourier transform of the series of length  $L$   $(v_1^{(k)}, v_2^{(k)}, \dots, v_K^{(k)}, 0, 0, \dots, 0)$  and then the sum  $\sum_{j=0}^{L-1}$  is replaced by  $\sum_{j=0}^{K-1}$ .

The forward filtering  $X\mathbf{v}$  can be rewritten in the Fourier space as

$$\hat{x}_\alpha \longrightarrow \hat{v}_{-\alpha}^{(k)} \hat{x}_\alpha = \hat{v}_\alpha^{(k)*} \hat{x}_\alpha \quad . \quad (7)$$

The above relation holds for the periodic series. Since the eigenvectors  $v_j^{(k)}$  are real valued, their Fourier transform satisfies  $\hat{v}_{-\alpha}^{(k)} = \hat{v}_\alpha^{(k)*}$ . In above equation, the vectors  $\hat{v}_\alpha^{(k)}$  are complex in general, and shift and distort the time series. The non-causal FIR filtering Eq. (7) corresponds to the SVD of the trajectory matrix in the conventional treatment of SSA. The last step of SSA is the anti-diagonal averaging (Hankelization) for the rank-one matrix  $X_k$  and this can be expressed as the causal FIR filtering

$$\hat{x}_\alpha \longrightarrow \hat{v}_\alpha^{(k)} \hat{x}_\alpha \quad . \quad (8)$$

Then the two-step filtering can be combined as a single real filtering

$$\hat{x}_\alpha \longrightarrow |\hat{v}_\alpha^{(k)}|^2 \hat{x}_\alpha \quad , \quad (9)$$

and does not distort the time series [1,9,10,12]. It should be important to note that the orthonormal and the completeness properties of the eigenvector  $\mathbf{v}^{(k)}$  lead to the following relations for the filters  $|\hat{v}_\alpha^{(k)}|^2$  [10,11]

$$\sum_{\alpha=0}^{L-1} |\hat{v}_\alpha^{(k)}|^2 / L = 1 \quad , \quad (10)$$

and

$$\sum_{k=1}^K |\hat{v}_\alpha^{(k)}|^2 / K = 1 \quad . \quad (11)$$

An arbitrary orthonormal  $K$ -dimensional vector set  $\{\mathbf{v}^{(k)}\}$  leads to the filters satisfying the relations Eqs. (10) and (11). In the SSA, the specific  $K$ -dimensional vectors  $\mathbf{v}$  are adopted which maximize the norm of the  $L$ -dimensional vector  $X\mathbf{v}$  preserving the orthonormality of the vectors  $\mathbf{v}$ , which leads to the eigenvalue equation Eq. (3). For the eigenvalues and the eigenvectors for the lagged-covariance matrix  $C$ , the following relation holds

$$\lambda_k = \mathbf{v}^{(k)T} X^T X \mathbf{v}^{(k)} = \sum_{\alpha} |\hat{v}_{\alpha}^{(k)}|^2 |\hat{x}_{\alpha}|^2 . \quad (12)$$

The SSA can be interpreted as an algorithm to determine the orthonormal  $K$ -dimensional vectors  $\mathbf{v}^{(k)}$  which maximize the quantity  $\sum_{\alpha} |\hat{v}_{\alpha}|^2 |\hat{x}_{\alpha}|^2$  preserving the orthonormality of the vectors  $\mathbf{v}$ . In this sense, the SSA can be understood as the principal component analysis applied to the trajectory matrix constructed from the lagged vectors. The orthogonality of the eigenvectors are preserved as the phase content for  $\hat{v}_{\alpha}^{(k)}$ , and then, it is lost for the filters  $|\hat{v}_{\alpha}^{(k)}|^2$ . Instead of losing the orthogonality, the filters  $|\hat{v}_{\alpha}^{(k)}|^2$  are real valued and lead no phase shift or distortion for the time series [1,9,10,12].

The SSA decomposition of the time series can simply be expressed as

$$\hat{x}_{\alpha} = (|\hat{v}_{\alpha}^{(1)}|^2 \hat{x}_{\alpha} + |\hat{v}_{\alpha}^{(2)}|^2 \hat{x}_{\alpha} + \dots + |\hat{v}_{\alpha}^{(K)}|^2 \hat{x}_{\alpha}) / K . \quad (13)$$

We denote the inverse Fourier transform of the component  $|\hat{v}_{\alpha}^{(k)}|^2 \hat{x}_{\alpha}$  as  $x_n^{(k)}$ . Then, the SSA decomposition of the time series is expressed as

$$x_n = x_n^{(1)} + x_n^{(2)} + \dots + x_n^{(K)} . \quad (14)$$

For later convenience, we denote the filters as

$$F_{\alpha}^{(k)} = |\hat{v}_{\alpha}^{(k)}|^2 , \quad (15)$$

and the Eqs. (10) and (11) can be rewritten as

$$\sum_{\alpha} F_{\alpha}^{(k)} / L = 1 , \quad (16)$$

and

$$\sum_k F_{\alpha}^{(k)} / K = 1 . \quad (17)$$

From the filtering viewpoint, the SSA algorithm can be seen as the optimal generation of the filter set satisfying the relations Eqs. (10) and (11), and the time-series decomposition can be done as in Eq.(13). From the completeness

relation Eq. (17), we can see that the decomposition of the power spectrum of the time series  $|\hat{x}_\alpha|^2$  is possible as

$$|\hat{x}_\alpha|^2 = \frac{1}{K} \sum_{k=1}^K F_\alpha^{(k)} |\hat{x}_\alpha|^2 . \quad (18)$$

Through the above decomposition of power spectrum, we can obtain the insights into the spectral structure of SSA decomposition. On the other hand, squared Fourier transform of the SSA decomposition  $x^{(k)}$  in Eq. (13) corresponds to  $[|\hat{v}_\alpha^{(k)}|^2 \hat{x}_\alpha]^2 = |\hat{v}_\alpha^{(i)}|^4 |\hat{x}_\alpha|^2$  and their sum does not reproduce the power spectrum of the time series. In the next section., we will show that the above decomposition can be applied to analyze the details of the SSA decomposition and it can be used to monitor the spectral structure of the decomposition which helps us for the proper choice of the window length parameter  $K$ .

### 3 Example

To illustrate how the spectral decomposition of the power spectrum described in the last section works, we consider the schematic time series and the daily currency exchange rate EUR/USD.

#### 3.1 Schematic model

First, we adopt the simple time series which is the superposition of the harmonic components with additive noise,

$$x_n = \sum_{i=1}^4 a_i \sin(2\pi n/b_i) + \epsilon_n , \quad (19)$$

where  $N = 300$  and  $\epsilon_n$  the Gaussian noise  $N(0, 1)$ . We adopt the parameters:  $a_1 = 1.44$ ,  $a_2 = 1.04$ ,  $a_3 = 1.04$ ,  $a_4 = 0.8$ ;  $b_1 = 29.9$ ,  $b_2 = 9.7$ ,  $b_3 = 8.9$  and  $b_4 = 4.9$ . The above time series Eq. (19) and its power spectrum are shown in Fig. 1. Corresponding to the harmonic components, there are four peaks A~D in the power spectrum. As seen, the second B and the third C peaks are close and have similar strengths. In such cases, these peaks are mixed with each other and are difficult to separate in SSA as is discussed in Ref. [3]. To see the dependence on the choice of the window

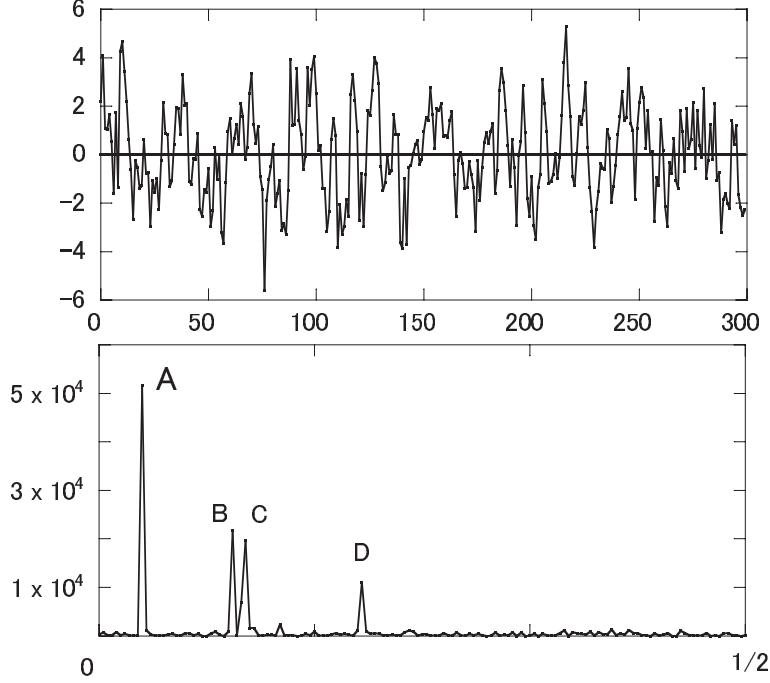


Figure 1: The upper figure is the time series in Eq. (19) and the lower is its power spectrum.

length, we adopt the values  $K = 40, 80$  and  $120$ . The larger 15 eigenvalues are shown in Fig. 2. It is known, for the case with larger window lengths, each harmonic component corresponds to the degenerate eigenvalue pair [3,4] and the corresponding eigenvectors are approximately the sinusoid with the relative phase shift  $\pi/2$  [3,13]. As seen, paired degeneracy of the eigenvalue becomes to be apparent for larger window lengths. For,  $K = 120$ , we can see clear four degenerate pairs. Since the filters are defined as the squared Fourier transform of the eigenvector of the lagged-covariance matrix as in Eq. (3), the Fourier transform  $\hat{v}_\alpha^{(k)}$  of the eigenvector for the degenerate pair differ only in phase and leads to approximately the same filters  $F^{(k)}$ . These filters extract the corresponding harmonic components from the time series. For all cases  $K = 40, 80$  and  $120$ , the peaks A and D are clearly filtered separately since these are isolated and the strengths are different from the other prominent spectral peaks. For the two peaks B and C, however, they are close with each other and the strengths are also similar. Though we can see the clear pairing of the eigenvalues for  $K = 120$  in Fig. 2, it is not clear to what



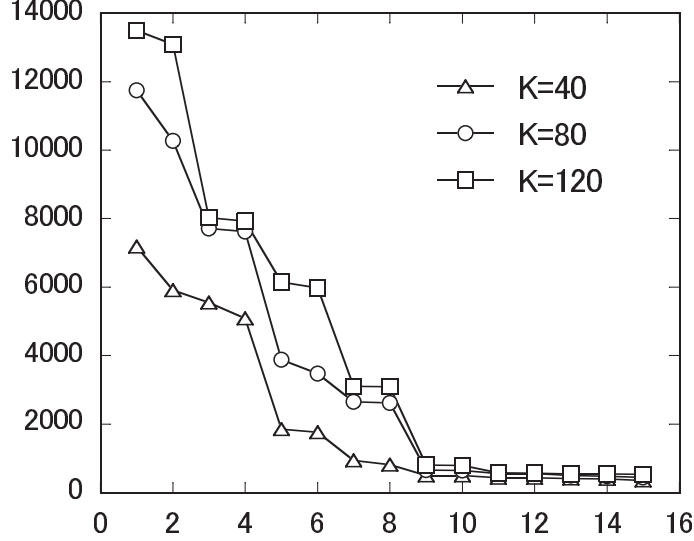


Figure 2: The larger 15 eigenvalues of the lagged-covariance matrix for the window lengths  $K = 40, 80$  and  $120$ , respectively. We can see that the eigenvalues are paired for larger  $K$ , corresponding to the harmonic components in the time series.

extent these peaks are separated in the SSA decomposition. To see this, we have shown in Fig. 3, the decomposition of the power spectrum in Eq. (18). Since the filter pairs  $(F^{(1)}, F^{(2)})$  and  $(F^{(7)}, F^{(8)})$  clearly extract the peaks A and D, respectively, these are not shown in the figures. The components  $F_{\alpha}^{(3)}|\hat{x}_{\alpha}|^2 \sim F_{\alpha}^{(6)}|\hat{x}_{\alpha}|^2$  are shown for the window lengths  $K = 40, 80$  and  $120$ , respectively. As seen, the filtered strengths are fragmented and the peak D component are mixed for smaller window length  $K = 40$ . For  $K = 80$ , the component of the peak D almost disappears and the four filters  $F_{\alpha}^{(3)} \sim F_{\alpha}^{(6)}$  extract both B and C components.

For larger window length  $K = 120$ , which is close to the half of the length of the time series, the filters  $F_{\alpha}^{(3)}$  and  $F_{\alpha}^{(4)}$  extract mainly the component C with small admixture of B component, while the filters  $F_{\alpha}^{(5)}$  and  $F_{\alpha}^{(6)}$  extract the component B with small admixture of C component. The vector components for the eigenvectors  $\mathbf{v}^{(1)}$  and  $\mathbf{v}^{(2)}$  are the sinusoid with  $\pi/2$  phase shift. The components of the paired eigenvectors  $(\mathbf{v}^{(3)}, \mathbf{v}^{(4)})$  or  $(\mathbf{v}^{(5)}, \mathbf{v}^{(6)})$  are the amplitude modulated harmonics with relative phase shift  $\pi/2$ , representing the mixing of B and C components. These structure is easy to see from the decomposition of the power spectrum. Instead of this, we can calculate the

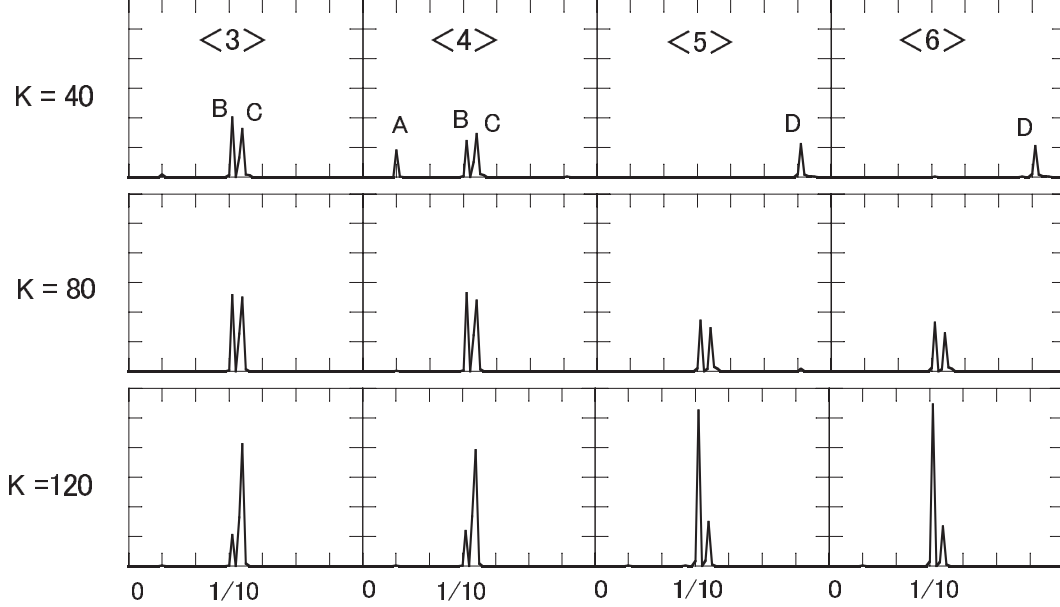


Figure 3: The spectral decompositions  $F_\alpha^{(3)}|\hat{x}_\alpha|^2 \sim F_\alpha^{(6)}|\hat{x}_\alpha|^2$  are shown for the window lengths  $K = 40, 80$  and  $120$ , respectively. From left to right, the figures represent the spectrum strengths  $F_\alpha^{(k)}|\hat{x}_\alpha|^2$  ( $k = 3 \sim 6$ ) denoted as  $\langle k \rangle$ . We take the same scale for all figures.

squared Fourier transform of the decomposition  $x^{(k)}$  as  $|F_\alpha^{(k)}\hat{x}_\alpha|^2$  but these quantities do not be summed up to reproduce the total power spectrum. The decomposition of the power spectrum has an advantage that the sum with respect to  $\alpha$  corresponds to the eigenvalues as in Eq. (12) and are summed up to reproduce the power spectrum as in Eq. (18).

In order to see the details of the distribution of the spectral decomposition, we have calculated the strength for individual spectral peaks as

$$I_i^{(k)} = \sum_{\alpha \in [\text{neighbor of the peak } i]} F_\alpha^{(k)}|\hat{x}_\alpha|^2 \quad i = A, B, C, D \quad , \quad (20)$$

and define the relative strengths  $R_i^{(k)}$  as

$$R_i^{(k)} = I_i^{(k)} / \sum_{j=1}^K I_i^{(j)} \quad . \quad (21)$$

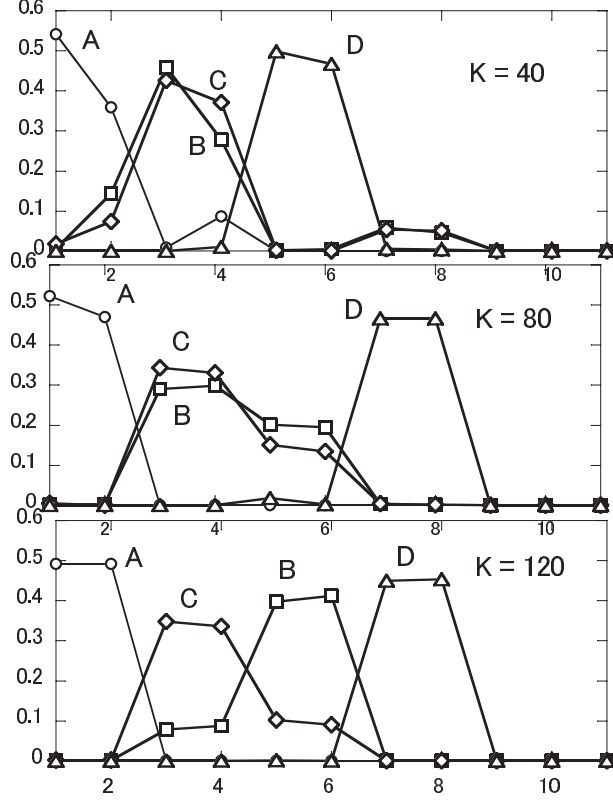


Figure 4: The relative strengths  $R_i^{(k)}$  ( $k = 1 \sim 11$ ) are shown for the spectral peaks  $i=A, B, C$  and  $D$  for the window lengths  $K = 40, 80$  and  $120$ , respectively.

The results of the calculated relative strengths  $R_i^{(k)}$  are shown in Fig. 4. As is expected, the strengths coming from the paired filters dominate for the isolated peaks A and D for almost all values  $K$  considered here. For  $K = 40$ , the peak D component is extracted with the filters  $F^{(5)}$  and  $F^{(6)}$ , while for larger window lengths, it is extracted with the filters  $F^{(7)}$  and  $F^{(8)}$ . For the peaks B and C, the strength distribution depends on the choice of the window length  $K$  as seen in the Fig. 4. For smaller values of  $K$ , two peaks B and C are considerably mixed with each other. For larger window length  $K$ , the peaks B and C are almost separated with the paired filters but still remains the small admixture of the components B and C. Thus, the SSA decomposition  $(x^{(1)}, x^{(2)})$  and  $(x^{(7)}, x^{(8)})$  are the pure harmonics with lower and higher frequencies, respectively, while the decompositions  $x^{(3)} \sim x^{(6)}$  are

the admixture of two harmonics.

For the case that the spectral peaks are well separated, the smaller window length is sufficient to separate the peaks of the spectrum and the results of the decomposition does not strongly dependent on the choice of the window length  $K$ . For the cases, where the spectral peaks are close or even overlap with each other, the Fourier peaks are mixed and also the SSA decomposition is sensitive to the choice of the window length parameter [3,4]. Thus, for the time series with the complex spectral structure, we should be careful for the choice of the window length parameter. For the proper SSA decomposition, it is helpful to monitor the spectral content of the decomposition and its dependence on the window length  $K$  by examining the power-spectrum decomposition as described here.

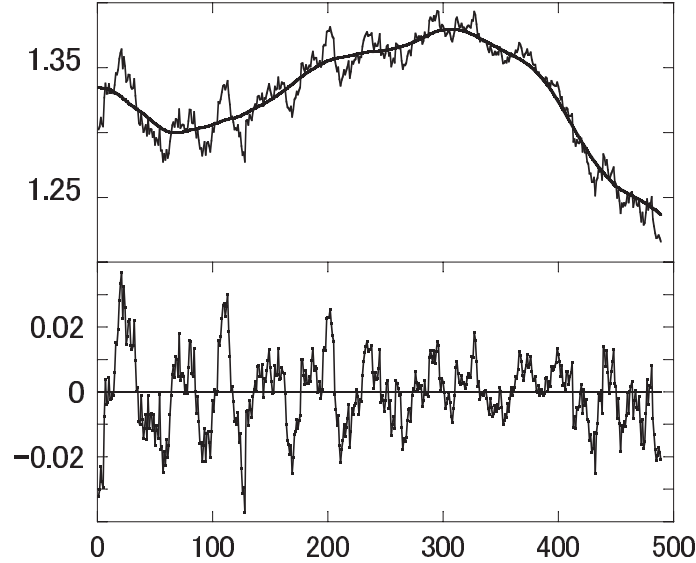


Figure 5: The daily currency exchange rate EUR/USD for the period 2013-2014. The number of data points is  $N = 489$ . The upper figure is the original daily rate and the first component of the SSA decomposition with the window length  $K=40$ . The lower graph is the residual time series.

### 3.2 Daily currency exchange rate EUR/USD

Next, we have applied the spectral decomposition for the analysis of the daily currency exchange rate Euro vs U.S. dollar for the period 2013-2014.

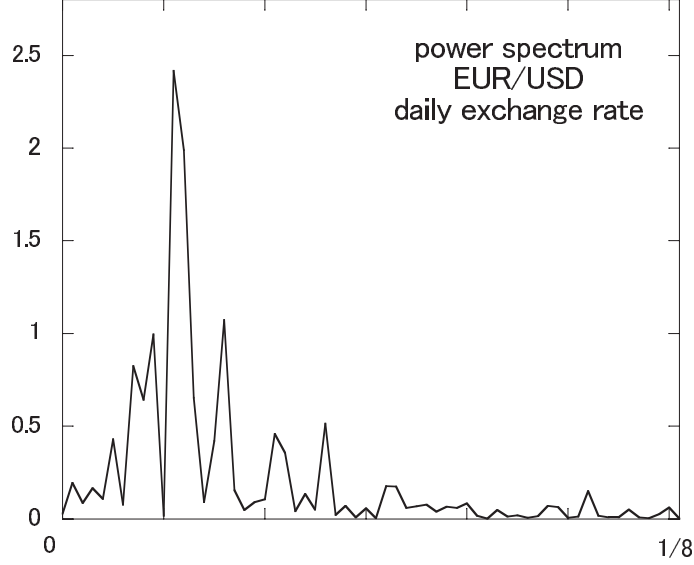


Figure 6: The power spectrum of the residual components of the currency exchange rate EUR/USD shown as the lower graph in Fig. 5.

The data consist of  $N=489$  points. The trend component of the series is separated by subtracting the first component of the SSA decomposition with the window length  $K = 40$ , which is called as the sequential SSA [3]. The original exchange rate and the first component of the SSA decomposition are shown in the upper graph in Fig. 5. The residual component is also shown as the lower graph, and its power spectrum is shown in Fig. 6. For the spectrum of the residual series, there are several peaks and are close with each other. Then, in order to see to what extent these harmonic components are separated, we have carried out the SSA decomposition with somewhat large window length parameters  $K = 140, 160$  and  $200$ . The eigenvalues  $\lambda_1 \sim \lambda_{13}$  are shown in Fig. 7.

In this case, the eigenvalues  $(\lambda_1, \lambda_2)$  and  $(\lambda_3, \lambda_4)$  are paired corresponding to the harmonic components in the series. The eigenvalues  $(\lambda_5, \lambda_6)$  seems to be paired but are close with the values for the pair  $(\lambda_7, \lambda_8)$  and these eigenvalues do not exhibit isolated pairing. There are five or six prominent peaks in the power spectrum and are distinguished by the attached indices A-F as in the upper graph of Fig. 8. In the lower three graphs of Fig. 8, we have shown the decomposition of the power spectrum for the case  $K = 140$ . Since the extracted components with pair of filters  $(F^{(1)}, F^{(2)})$ ,  $(F^{(3)}, F^{(4)})$

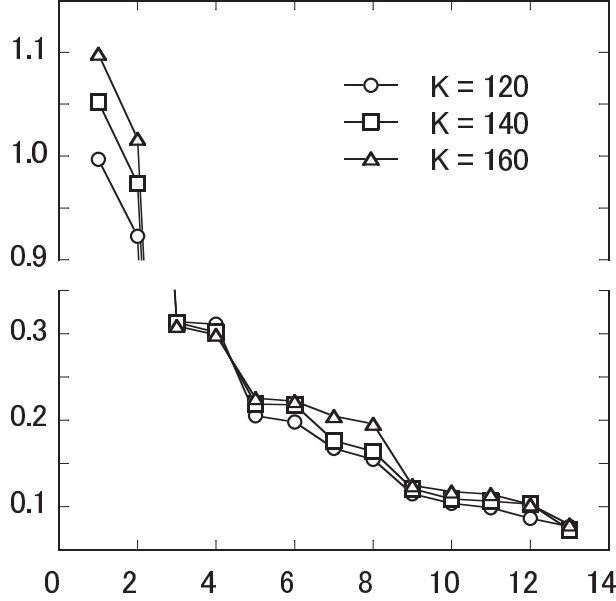


Figure 7: The larger 13 eigenvalues of the lagged-covariance matrix for the residual components with the window lengths  $K=120$ , 160 and 200, respectively. The first and the second eigenvalues  $\lambda_1$  and  $\lambda_2$  dominate.

and  $(F^{(5)}, F^{(6)})$  are almost the same and we cannot distinguish them in the figure, we have shown the sum of these components  $(F_{\alpha}^{(1)} + F_{\alpha}^{(2)})|\hat{x}_{\alpha}|^2$ ,  $(F_{\alpha}^{(3)} + F_{\alpha}^{(4)})|\hat{x}_{\alpha}|^2$  and  $(F_{\alpha}^{(5)} + F_{\alpha}^{(6)})|\hat{x}_{\alpha}|^2$  in these figures. The pair of filters  $(F^{(1)}, F^{(2)})$  corresponds to the largest peaks C, while the  $(F^{(3)}, F^{(4)})$  and  $(F^{(5)}, F^{(6)})$  pairs extract the peak D and E components, respectively. The broader peak B consists of the components  $\lambda_7$  and  $\lambda_8$ , but in this case, there is moderate admixture of dominant C component. The peaks A and B are close with each other and the strengths are the mixture of  $\lambda_9$  and  $\lambda_{10}$  components. In order to see the dependence on the choice of the window length  $K$ , we have calculated the relative strengths in Eq. (21) for the six peaks A-F and the results are shown in Fig. 9 for  $K = 120$ , 140 and 160, respectively. For the dominant harmonic components C, D and E, the results do not strongly depend on the choice of window length  $K$ . For the weak components A, B and F, the details of the filtering are considerably dependent on the choice of the window length  $K$ . For the separation of the harmonic components, window length around  $K = 140$  seems to be appropriate.

In the SSA, the filters  $F_{\alpha}^{(k)}$  are optimally constructed from the time series.

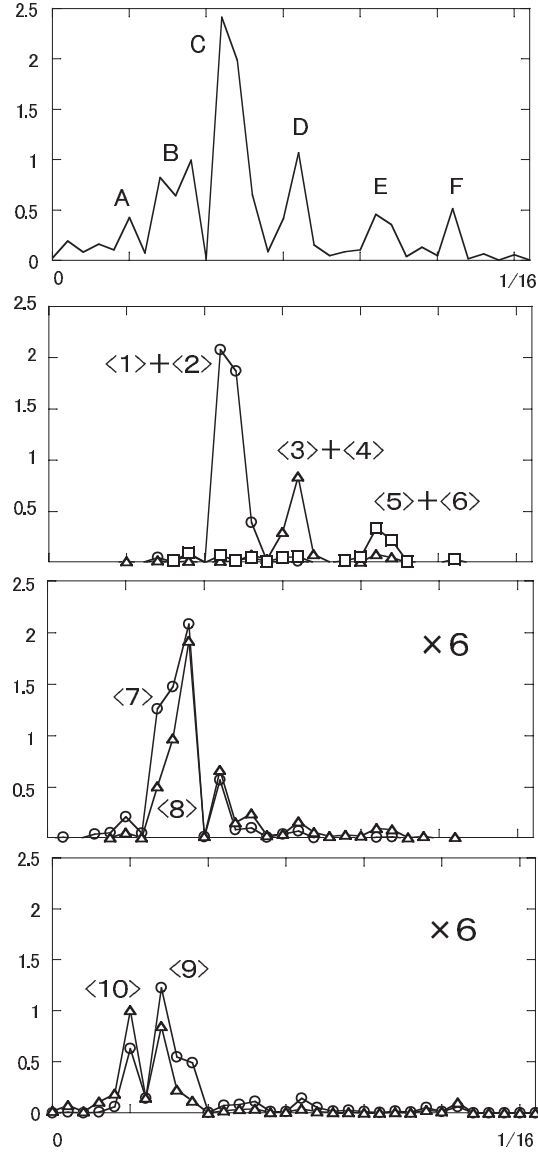


Figure 8: The spectral decompositions  $F_{\alpha}^{(1)}|\hat{x}_{\alpha}|^2 \sim F_{\alpha}^{(10)}|\hat{x}_{\alpha}|^2$  are shown for the window length  $K = 140$ . The individual components  $F_{\alpha}^{(k)}|\hat{x}_{\alpha}|^2$  are abbreviated as  $\langle k \rangle$ . For the lower two graphs, a factor 6 is multiplied.

We can say that the SSA algorithm is the soft partitioning of the Fourier spectrum with the normalized filters. Roughly speaking, the average widths of these filters in Fourier space gradually decreases with the increase of the

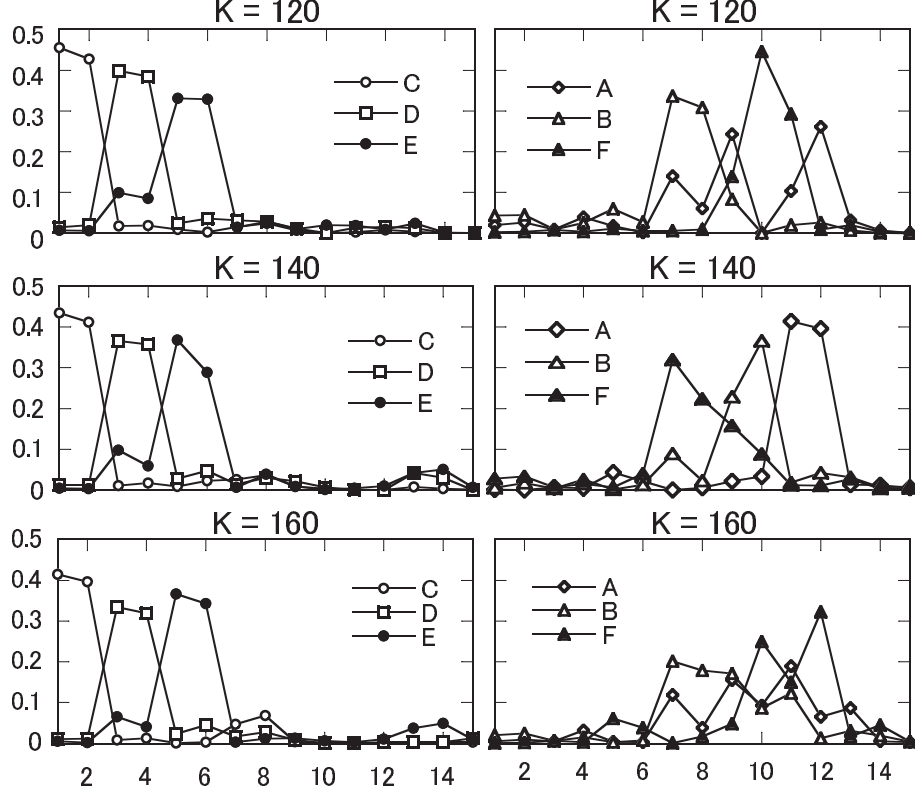


Figure 9: The relative strengths  $R_i^{(k)}$  ( $k = 1 \sim 15$ ) are shown for the peaks  $i=A \sim F$  with the window lengths  $K = 120, 140$  and  $160$ .

window length  $K$  extracting the finer structure of the spectrum, but the structure of these filters are strongly dependent on the power spectrum of the time series itself. Generally, for dominant eigenvalues, corresponding filter  $F_\alpha^{(k)}$  has a simple structure of prominent peak and oscillating tails similar to the squared sinc function. For non-dominant eigenvalues, however, the filters have complicated and broader function of  $\alpha$ . Thus, the SSA decomposition is considerably dependent on the choice of window length and it is difficult to find the general criterion for the proper choice of the window length  $K$ . Because of these, the spectrum decomposition with the filters  $F^{(k)}$  are useful to monitor the SSA decomposition in detail and adjust the window length in a specific application of the time series.



## 4 Summary and conclusions

In this paper, based on the filtering interpretation of SSA, the decomposition of the power spectrum for the original time series is shown to be possible and this helps us for obtaining the detailed insights into the SSA decomposition and it also helps us to choose the proper window length parameter  $K$  for a specific application. In the filtering interpretation, the SSA algorithm has been seen as the generation of normalized and complete set of adaptive filters and the SSA decomposition of the time series can be done through the operation of these filters to the original data. The SSA decomposition is the soft and the optimal partitioning of the Fourier space through the adaptively constructed real-valued filters. By using the completeness property of the eigenvectors of the lagged-covariance matrix, not only the time series itself but also the power spectrum can be decomposed into *additive*  $K$  components with the filters. The decomposition of the spectrum clearly shows the detailed structure of the SSA decomposition. Roughly speaking, the widths of the filters  $F_{\alpha}^{(k)}$  slowly decrease with the increase of window length  $K$ . This suggests that the finer structure of the spectrum of the time series can be separated with larger window length. But the structure of the filters are fairly complex and the dependence on the window length is not so simple as to set any general criterion for the proper choice of the window length  $K$ . Besides eigentriples of the SSA decomposition (eigenvalues, right and left singular values of the trajectory matrix), the decomposition of the power spectrum with the filters is useful for monitoring how the SSA algorithm works for the time-series decomposition and for the proper choice of the window length  $K$ .

### Acknowledgment

This work is supported by Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science; Grant Numbers (24650150) and (26330237).

## References

- [1] Bozzo, E., Carniel, R. and Fasino, D. (2010). Relationship between singular spectrum analysis and fourier analysis: theory and application

to the monitoring of volcanic activity. *Comput. Math. Appl.*, **60**:812–820.

- [2] Elsner, J. B. and Tsonis, A. A. (1996). *Singular Spectrum Analysis. A new tool in time series analysis*. Plenum Press.
- [3] Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001). Analysis of time series structure: SSA and related techniques. *Chapman and Hall/CRC*.
- [4] Golyandina, N. and Zhigljavsky, A., (2013). Singular spectrum analysis for time series *Springer*.
- [5] Ghil, M., Allen, M. R., Dettinger, M.D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F. and Yiou, P. (2002). Advanced spectral methods for climatic time series. *Rev. Geophys.*, **40-1**: 1–41.
- [6] Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *J. Data Sci.*, **5**: 239–257.
- [7] Hassani, H. and Zhigljavsky, A. (2009). Singular spectrum analysis: Methodology and application to economic data, *J. Syst. Sci. Complex. (JSSC)*, **22**: 372–394.
- [8] Harris, T. J. and Yuan, H. (2010). Filtering and frequency interpretations of singular spectrum analysis. *Physica D*, **239**: 1958–1967, and references therein.
- [9] Kormylo, J. and Jain, V. (1974). Two-pass recursive digital filter with zero phase shift, *IEEE Trans. Acoustics, Speech and Signal Process.*, **22**: 384–387.
- [10] Kume, K. (2012). Interpretation of singular spectrum analysis as complete eigenfilter decomposition. *Advances in Adaptive Data Analysis*, **4-4**: 1250023 1–18.
- [11] Kume, K. and Nose-Togawa, N. (2014). Multidimensional extension of singular spectrum analysis based on filtering interpretation. *Advances in Adaptive Data Analysis*, **6-1**: 1450005 1–17.

- [12] Tomé, A. M., Teixeira, A. R. , Figueiredo, N., Santos, I. M., Georgieva, P. and Lang, E. W. (2010). SSA of biomedical signals: A linear invariant system approach, *Statistics and Its Interface*,**3**: 345–355.
- [13] Vautard, R. and Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, *Physica D* **35** 395-424.